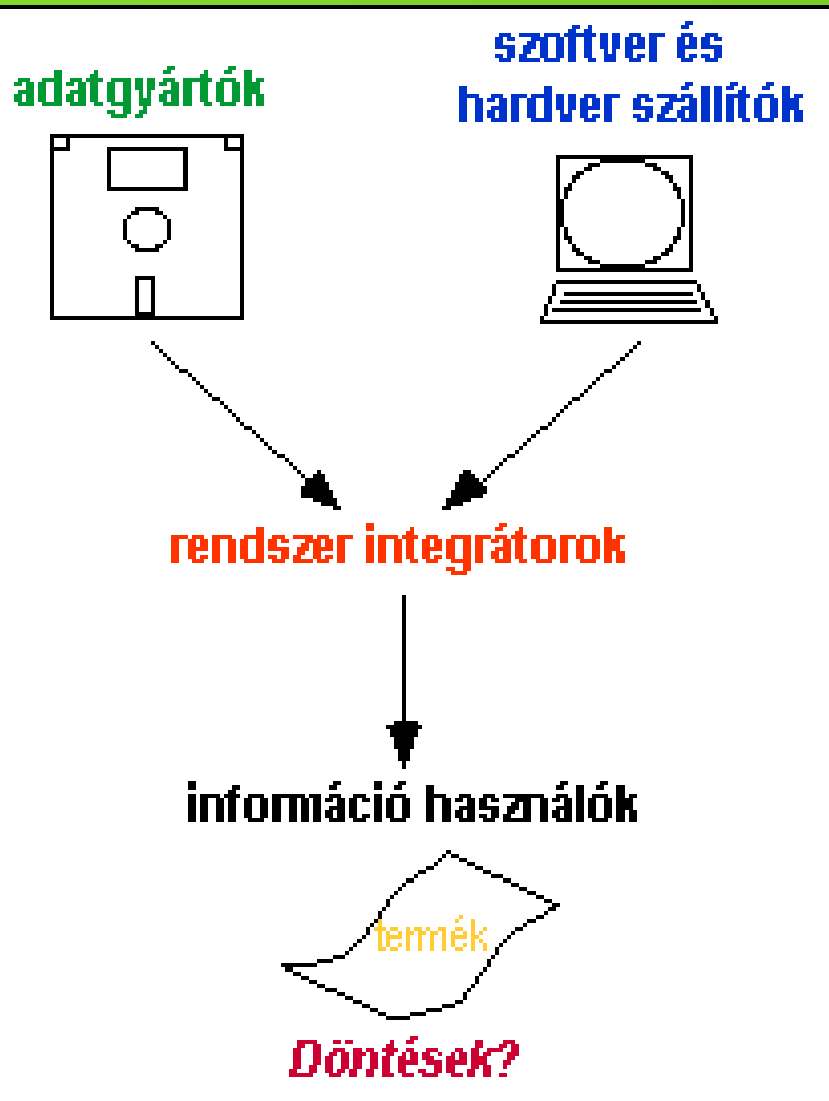


# **Adatbázisok és GIS műveletek pontossága**

# A bizonytalansági vita résztvevői



A digitális adatoktól és a létrehozott termékektől is elvárható hogy adott pontossági jellemzőkkel rendelkezzen.

A pontosság (helyzeti, időbeli vagy attribútum) és az élesség közötti különbség.

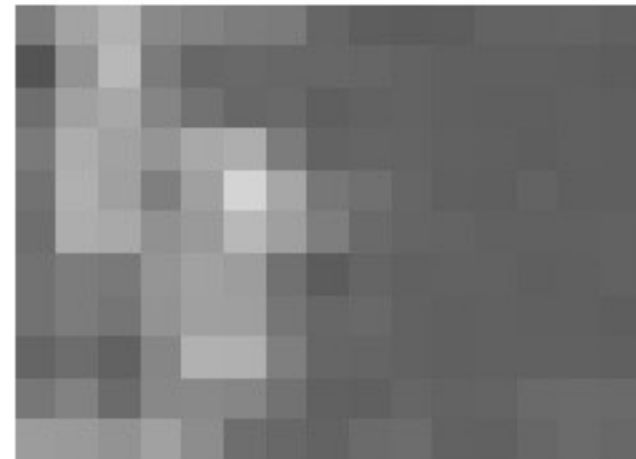
A metaadatok szabványosítása lehetővé tette, hogy a digitális adatok is tanúsított pontossági jellemzőkkel rendelkezzenek.

**Az élesség az értékes számjegyek száma, ha a pontosság 1m akkor a milliméteres élességnek nincs értelme. Ha azonban a pontosság 1 milliméter akkor az 1 méteres élesség degradálja a pontosságot!**

**A felbontás hasonló az élességhez. Ha az eredeti felvevő berendezés pixel mérete 40 m. volt, a több pixelen való megjelenítés nem fogja javítani a kép információ tartalmát**



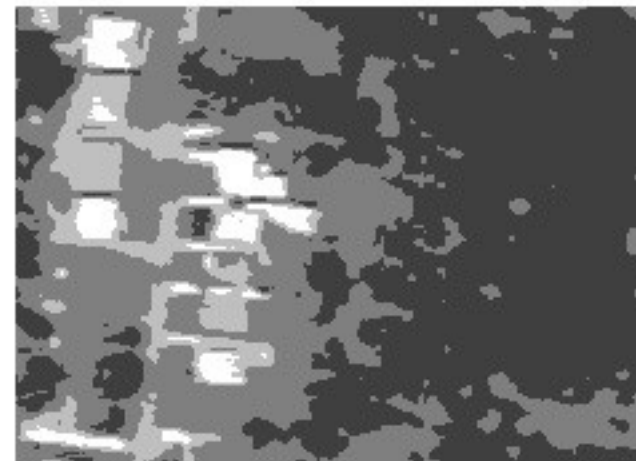
(a)



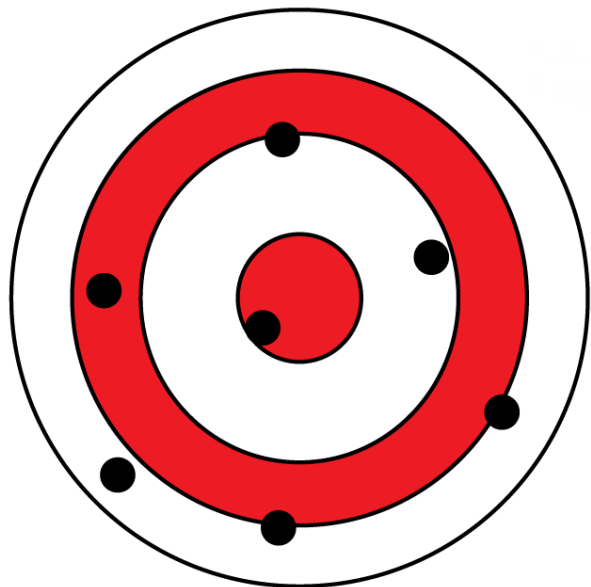
(b)



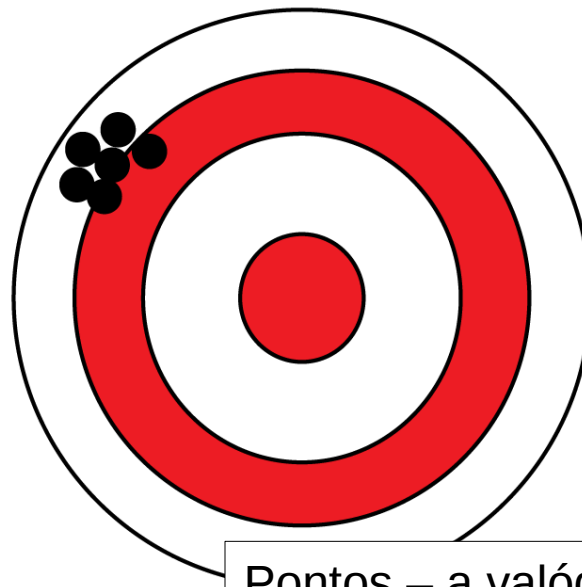
(c)



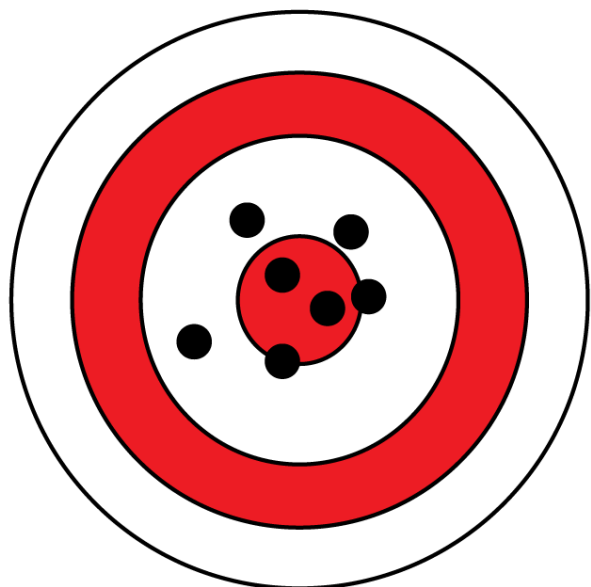
(d)



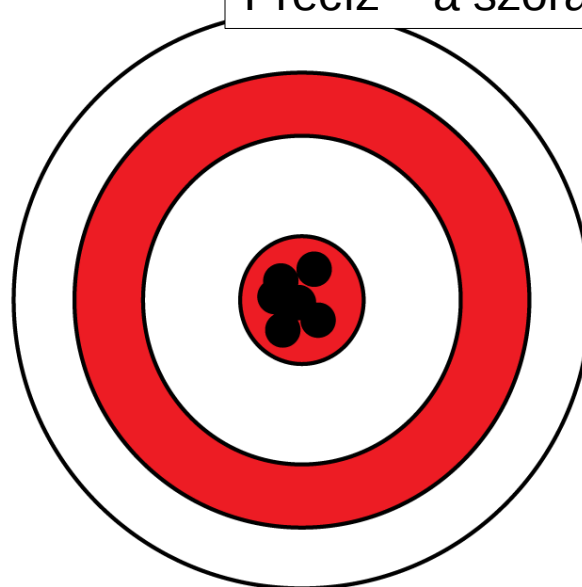
Pontatlan és  
nem precíz



Pontatlan, de  
precíz



Pontos, de  
nem precíz



Pontos és  
precíz

Pontos – a valódi értékkel egyezik  
Precíz – a szórás kicsi

# Az adatminőség komponensei

- A helyzeti pontosság
- Attribútum pontosság (vagy mérési hiba mint pld. a DEM-ben, vagy osztályozási hiba)
- Logikai konzisztencia (a topológiai kapcsolatok helyessége: zárt poligonok, minden poligonnak csak egy címkéje van)
- Teljesség – a kérdéses témában az adott területen található valamennyi objektum adatbázisban van-e
- Eredet az állomány létrehozásának története, a technológiák, pontossági mérőszámok, időpontok, végrehajtást végző személyek ismertetése.

# Helyzeti hibák eredete

Adatnyerés – alappontok, részletmérés, fotogrammetriai kiértékelés hibái;

•(papír) térképek szerkesztési elvei miatt:

•Generalizálás - vonalsimítás

•Összevonás

•Eltolt ábrázolás

•Számítási – feldolgozási hibák

•Hibás interpretálás

•Kerekítési hibák

•Vektor-raszter, raszter-vektor átalakítás

# A Cohen-Kappa egyezési tényező

Más a helyzet az osztályozás megbízhatóságának eldöntésénél mint a helyzeti pontosság meghatározásánál, mivel az osztály nominális érték és nem helyezhető el a számegyenesen.

Erre az esetre egy *Cohen* nevű pszichológus 1960-ban kidolgozott statisztikai mutatóját a *kappa egyezési tényezőt* használják, melyet *Congalton* és *Mead* vezettek be 1983-ban a távérzékelésbe.

A módszer lényegét lássuk egy példán.

Képzeljük el, hogy négy osztályunk van: A, B, C és D. Vizsgáljuk meg a helyszínen  $N=60$  esetben hogy a kérdéses pont melyik osztályba tartozik és nézzük meg ugyanezen pontok hovatartozását az osztályozott távérzékelési mintában.

Osztály a távérzékelte adatokból	Osztály a terepen				
	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>összeg</b>
<b>A</b>	<b>5</b>	<b>4</b>	<b>2</b>	<b>3</b>	<b>14</b>
<b>B</b>	<b>1</b>	<b>8</b>	<b>1</b>	<b>3</b>	<b>13</b>
<b>C</b>	<b>5</b>	<b>2</b>	<b>4</b>	<b>3</b>	<b>14</b>
<b>D</b>	<b>6</b>	<b>5</b>	<b>1</b>	<b>7</b>	<b>19</b>
<b>összeg</b>	<b>17</b>	<b>19</b>	<b>8</b>	<b>16</b>	<b>60</b>



Belátható, hogy hibátlan osztályozás esetén csak a táblázat átlójában szerepelnének zérustól eltérő számok.

Jelöljük  $d$ -vel az **egyező egységek arányát** (azaz az átlóban lévő elemek összegének a viszonyát az összes méréshez):

$$d = \frac{(5+8+4+7)}{60} = 0.40$$

majd határozzuk meg a **véletlen egyezések arányát** mint a sor és oszlopösszegek szorzatának hányadosát az összes mérési eredmény négyzetével és összegezzük ezt valamennyi osztályra:

$$q = \frac{14 \cdot 17}{60 \cdot 60} + \frac{13 \cdot 19}{60 \cdot 60} + \frac{14 \cdot 8}{60 \cdot 60} + \frac{19 \cdot 16}{60 \cdot 60} = 0.25$$

Ezek után a **kappa egyezési** tényezőt a következő kifejezésből nyerjük:

$$k = \frac{d - q}{1 - q} = \frac{0.40 - 0.25}{1 - 0.25} = 0.20$$

Hibátlan osztályba sorolás esetén  $\kappa=1$ .

A KAPPA 0-tól 1-ig terjed. A 0 az egyezés teljes esetlegességére utal, míg az 1 a tökéletes valódi egyezésre. Ha az érték pl. 0.145 ezt úgy tekinthetjük, hogy az egyezés 14 százalékkal jobb annál, mintha csak teljesen esetleges lenne.

# Hibaterjedés

A több réteggel végrehajtott műveletek eredményeképpen

- A leggyöngébb réteg pontossága határozza meg a hibát ha a műveletben AND operátort használunk
- Az OR művelet alkalmazása esetén az eredmény pontossága jobb mint a leggyengébb láncszemé

Az érzékenység elemzés segítségével tudjuk meghatározni, ahhoz, hogy az eredmény megadott pontosságú legyen milyen pontossága kell, hogy legyen az egyes rétegeknek.

- Egységgel megváltoztatjuk a numerikus réteg cella értékeit és megnézzük, hogy miként hat az eredményre
- Nem numerikus rétegen meg kell vizsgálni, hogy melyik változás rontja el legjobban az eredményt
- A súlyok változására is el kell végezni ezeket a műveleteket

# Fuzzy matematika (Zadeh 1965)

Térbeli bizonytalanság kezelése

Szubjektív tényezők figyelembevételét teszi lehetővé

0/1 értékek helyett több kategória, vagy folytonos átmenet, az emberi logikához jobban illeszkedik

Membership függvény:  $\mu(l) \in [0,1]$

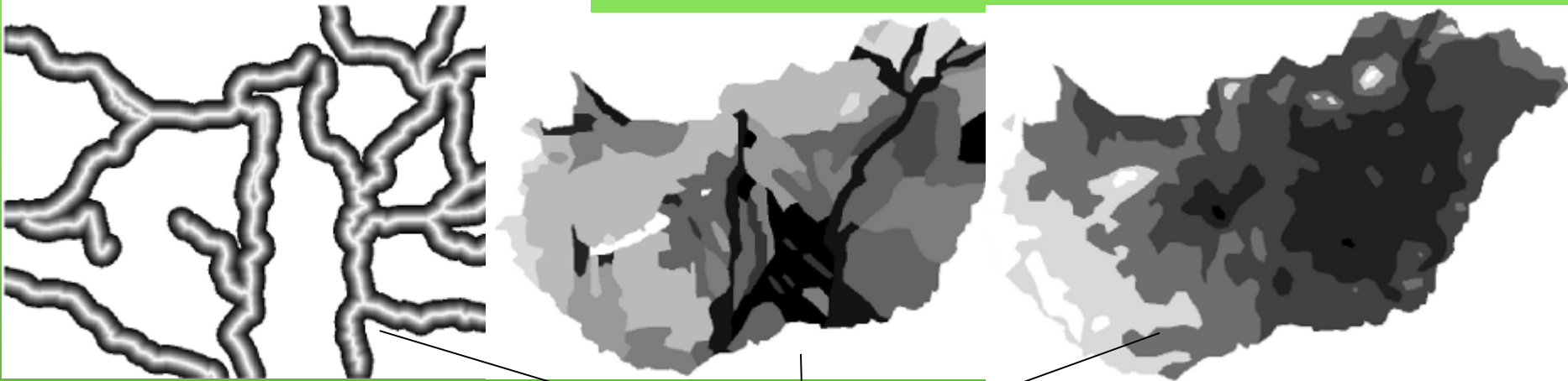
Halmaz műveletek: Komplement  $\forall l \in L : \mu_3(l) = 1 - \mu_1(l)$

Unió  $\forall l \in L : \mu_3(l) = \max\{\mu_1(l), \mu_2(l)\}$

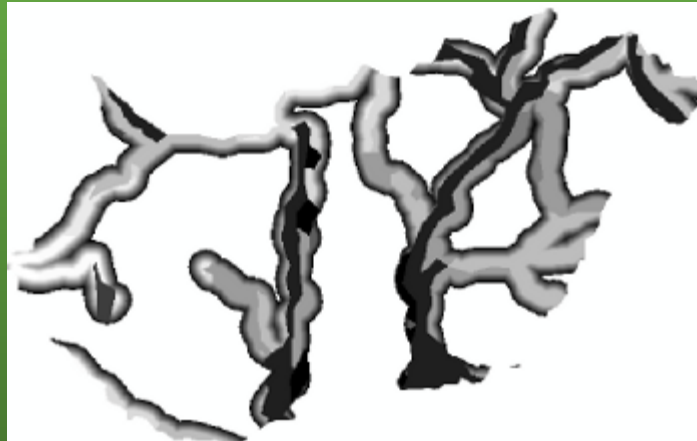
Metszet  $\forall l \in L : \mu_3(l) = \min\{\mu_1(l), \mu_2(l)\}$

Raszteres adatokkal egyszerűen megvalósítható,  
GRASS-ban megvalósítható

# Fuzzy elemzés



Min(.....)



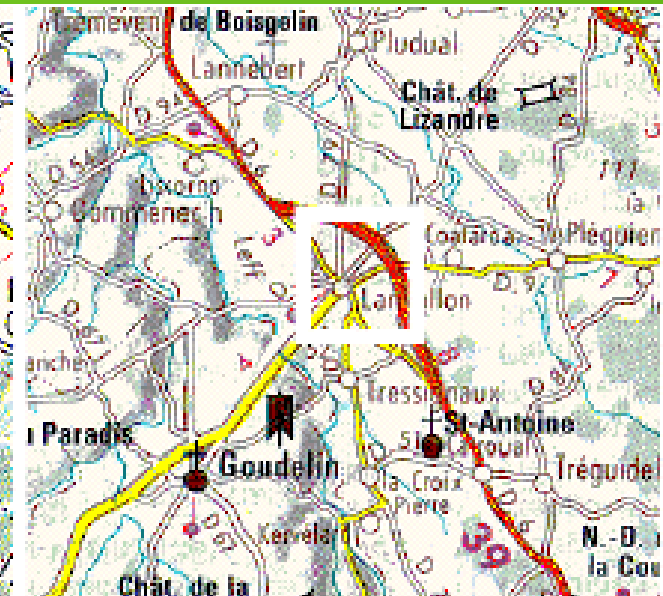
# A generalizálás



1:25 000

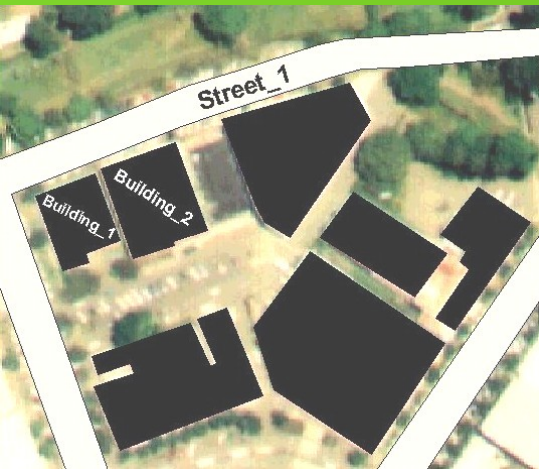


1:100 000

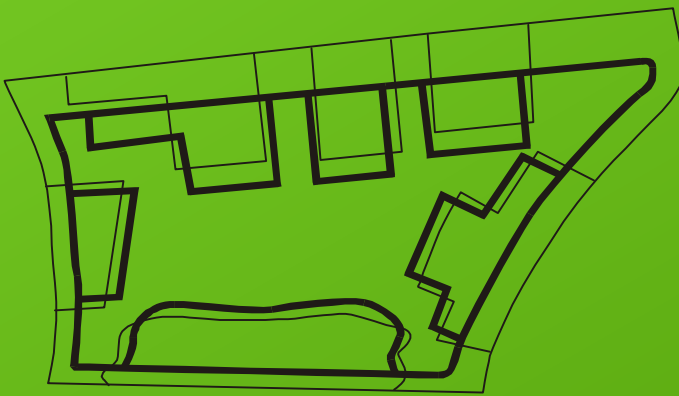
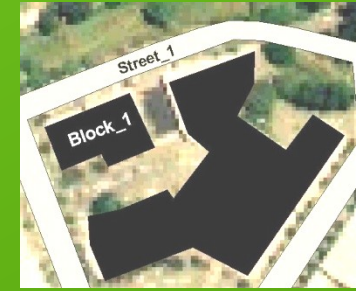


1:250 000

Ha nagyfelbontású adatbázist kis felbontásban akarunk megjeleníteni: generalizálni kell, mivel a szem úgysem tudja követni a részleteket, ki kell hangsúlyozni a lényegét, a méretarányon felüli szimbólumos ábrázolásból eredő takarásokat meg kell szüntetni.



Példák különböző szintű összevonásra



Példa az eltolások következtében alkalmazott affin torzításra

# A generalizálást több hierarchia szinten kell megvalósítani

*Az első szint* azoknak az objektumoknak a megválasztása, amelyeknek a kérdéses térképen egyáltalán szerepelnie kell.

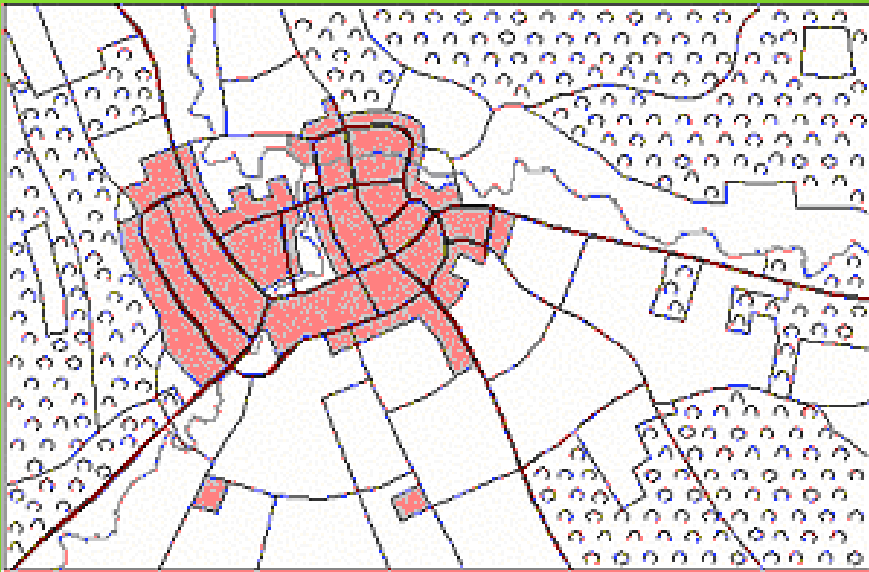
*A második szinten* az objektumok összevonása történik összefoglaló objektumokká pl. a különálló házakból tömböket alakítunk ki.

*A következő hierarchia szinten* a megmaradt egyszerű vagy összetett objektumokat vizsgáljuk meg abból a szempontból, hogy valós méreteik meghaladják-e a méretarányhoz rendelt minimumot, ha igen úgy a következő lépésben egyszerűsítünk a (kör)vonalaikon, ha nem adott méretű **egyezményes jellel helyettesítjük**.

A következő lépésben egyszerűsítjük a nem egyezményes jellel ábrázolt, megmaradt vonalas objektumok határvonalait, ezt nevezzük vonalgeneralizálásnak.

**Az első két szint: MODELL generalizálás,  
a többi: Kartográfiai generalizálás**

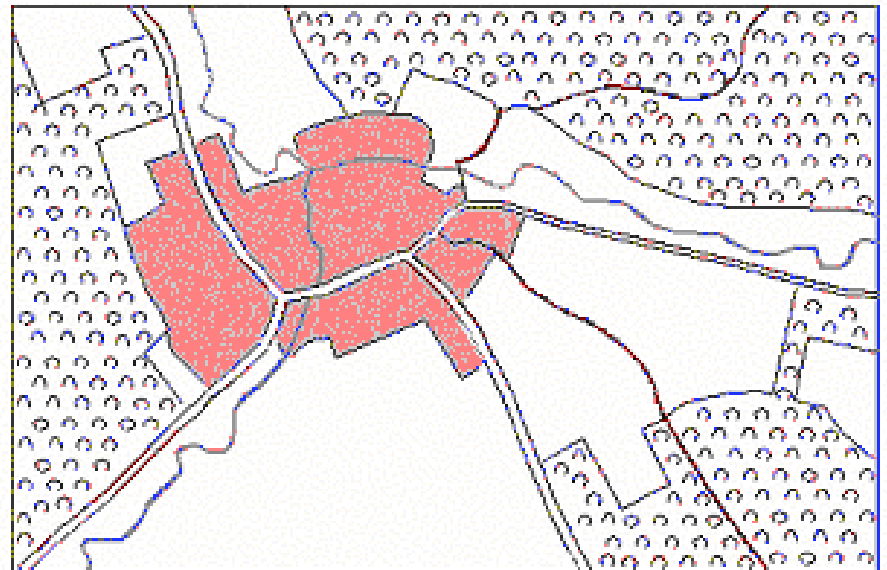
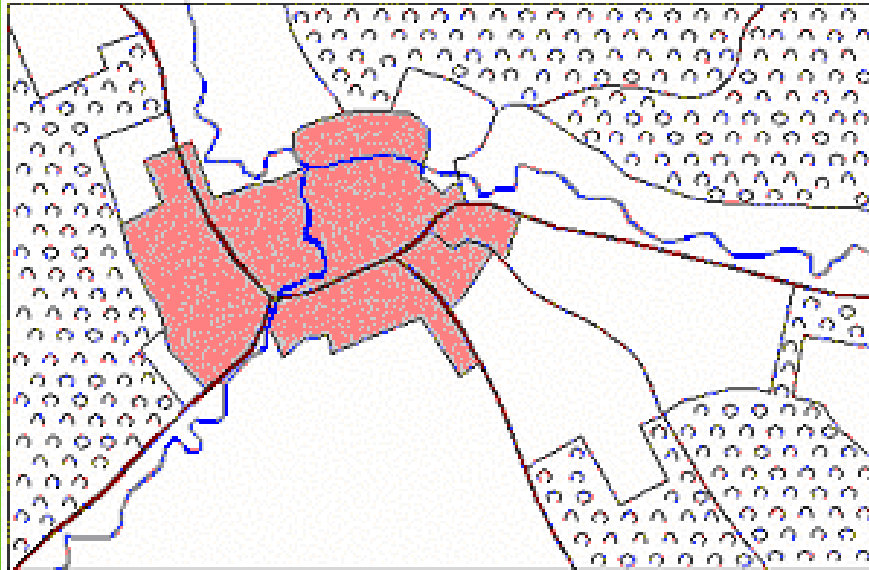
*Mivel az egyezményes jelek általában nagyobbak mint a tényleges tárgyak méretarányban kifejezett képei ezért két közeli egyezményes jel gyakran fedheti egymást. Ezt elkerülendő, a kevésbé fontos objektum szimbólumát a generalizálás során el kell tolni.*



**Modell generalizálás**



**Kartográfiai generalizálás**

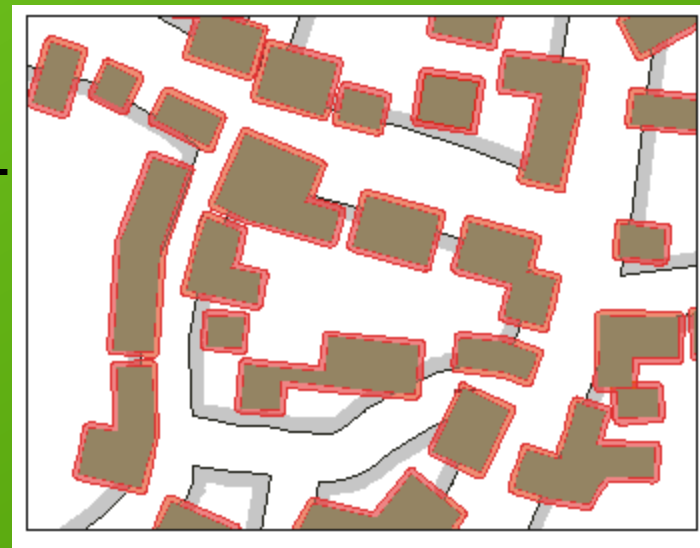




A bal oldalon a kataszteri térkép épületei, a jobb-  
oldalon az épületek összevonása a topo-  
gráfiai méretarányban a **CHANGE** program-  
mal



Balra az utak szimbólumai és védőövezetei, jobbra az összes szimbólum védőövezettel





A konfliktus feloldása után, a bal oldalon a „barát” objektumok vonzását nem vettük figyelembe, a jobboldalon viszont igen (a házak rásimulnak az útra)



Output a topográfiai térképre

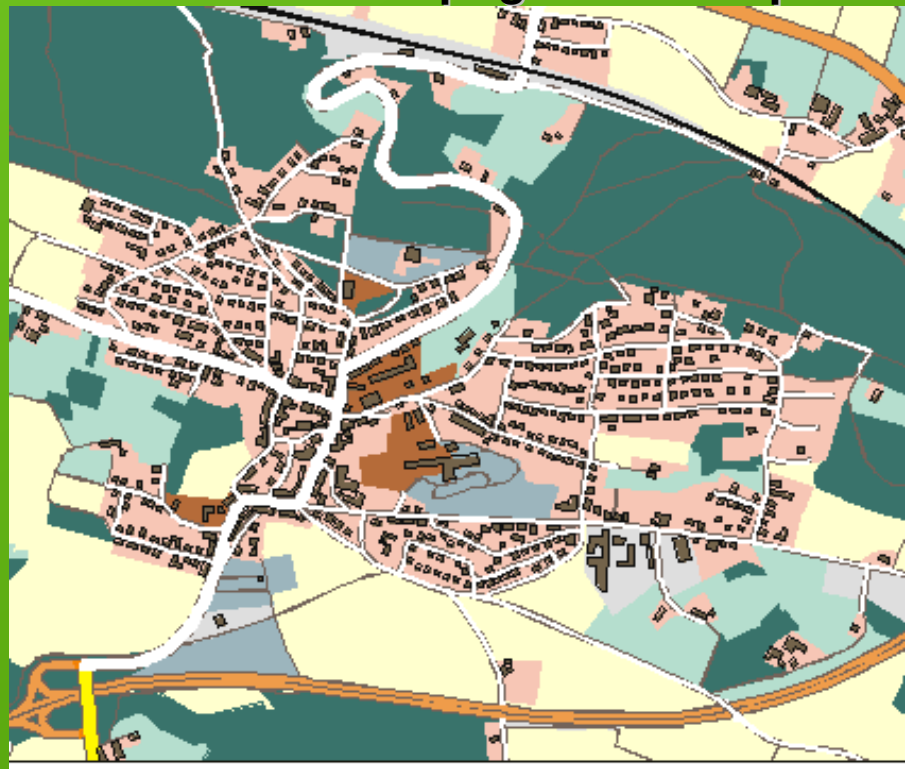
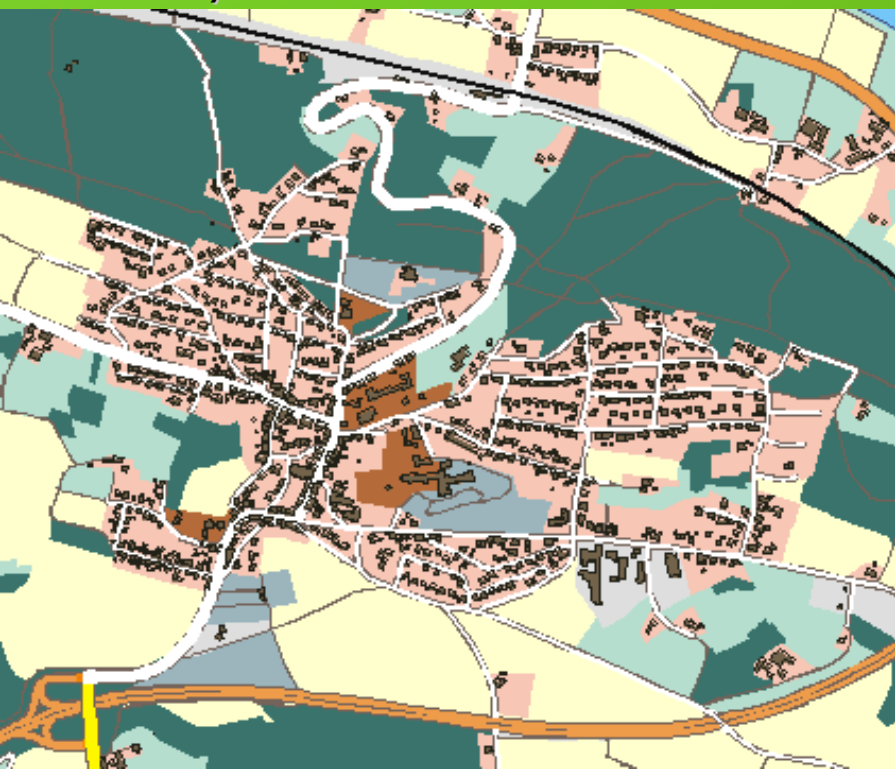




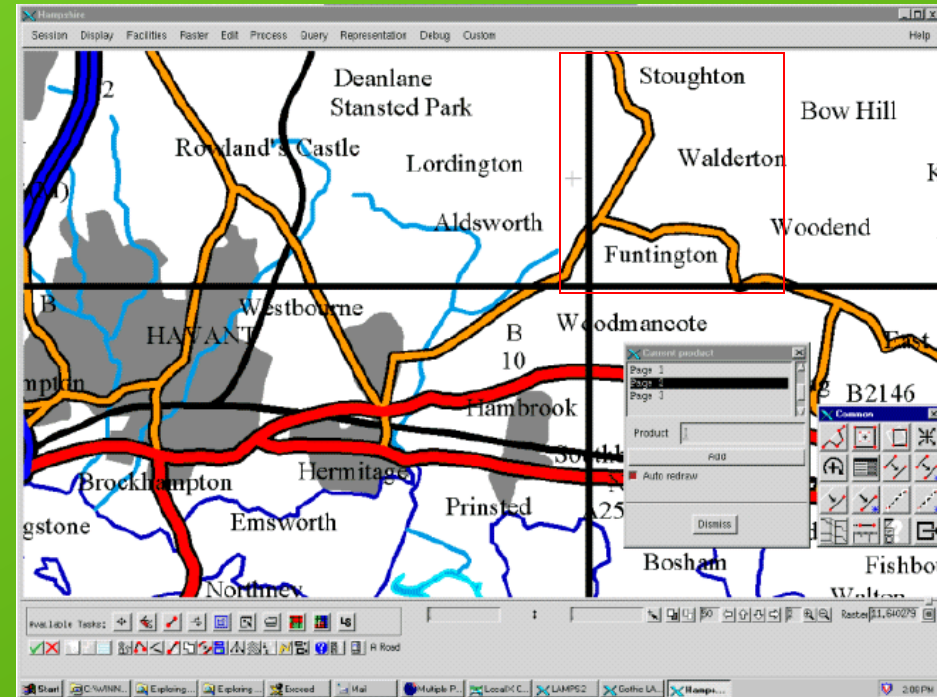
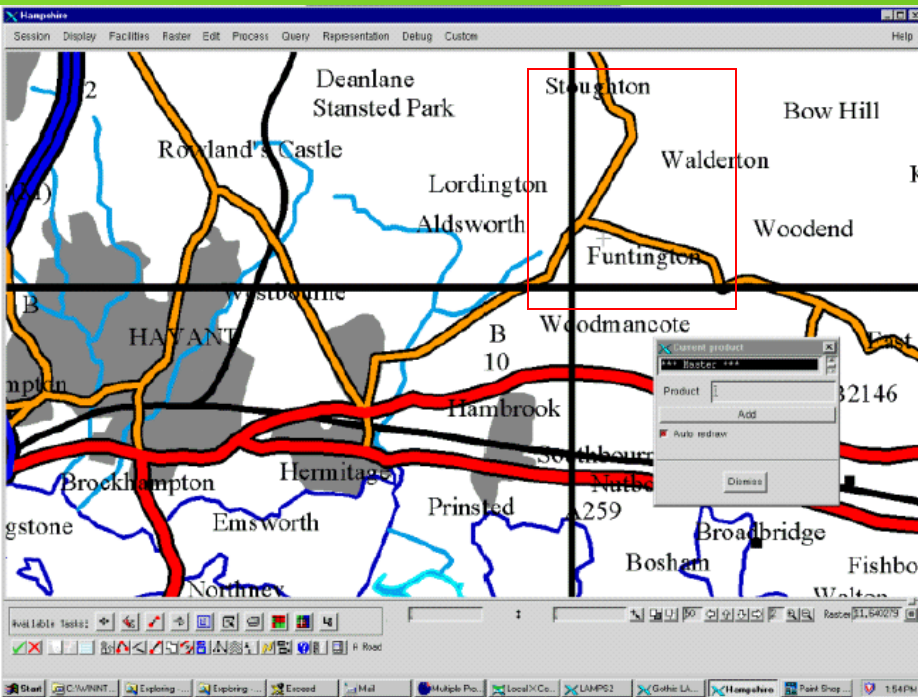
**Az eredeti, kataszteri  
tartalom és a  
generalizált outputok  
összehasonlítása**



**Lent, a kataszteri ház tematikával kiegészített 1:25000-es topográfiai térkép**



**Objektum-orientált adatbázisban több geometria is tartozhat ugyanahhoz az adathoz, a piros téglalapban lévő falnevek eredeti tárolása a baloldali képen látható, de ugyanezek az adatok megjeleníthetők a jobboldali alternatív geometriában is.**



**Ugyanezen az elven nyugszik a több méretarányú (multi-scale) megjelenítés is! Kísérletezzünk az openstreetmaps-el vagy a Google maps-el**

# GIS és a generalizálás

Eltérő méretarányokban jelenítjük meg a térképeinket (zoom)

Digitális térkép méretaránya?

Az eredeti adatokat nem szabad degradálni (generalizálni) –  
az elemzések eredményét torzítaná

Több különböző méretarányhoz tartozó megjelenítést nehéz  
karbantartani

Röptében történő megoldások

pl. vonalsimítás a megjelenítés gyorsítására,

méretarány függő megjelenítés/elrejtés,

méret (terület, hossz) alapján történő megjelenítés/elrejtés

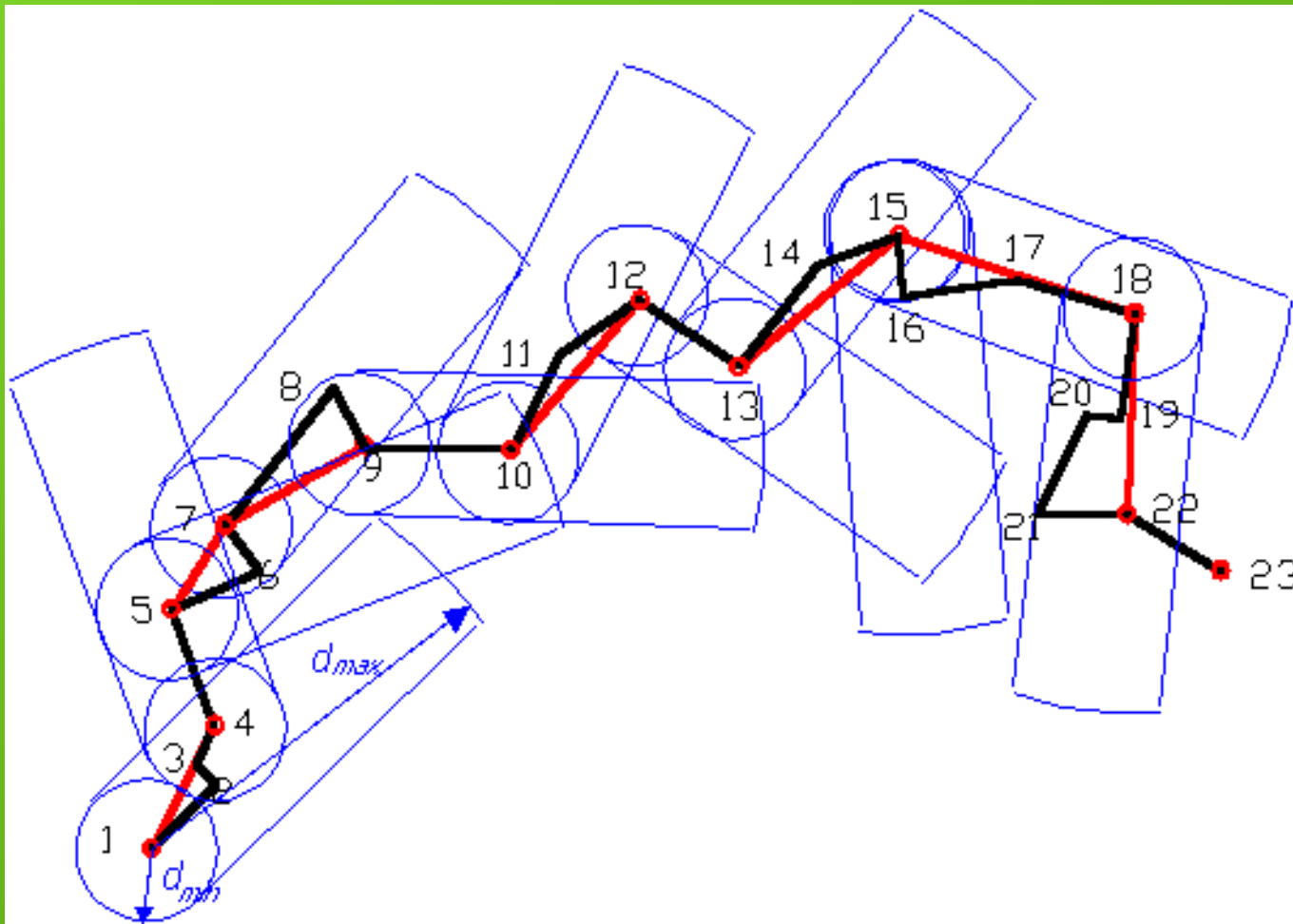
egymást takaró címkék elhagyása, dinamikus elhelyezés,

közeli pontok összevonása a megjelenítésben,

egybeeső pontok eltolása a megjelenítésben

## VONALGENERALIZÁLÁS

**Opheim** egyszerűsítési algoritmusából indul ki több más algoritmussal egyetemben, hogy azokat a pontokat lehet elhagyni, melyek egy-egy vonaldarabra illesztett párhuzamos egyenesek által létrehozott folyosón belül helyezkednek el. Ezen kívül még olyan kikötést is tesz, hogy a pontok nem lehetnek egy távolságnál közelebb illetve egy távolságnál ritkábban



# A *Douglas* féle egyszerűsítési algoritmus *globális eljárás*, tehát a *vonalat teljes terjedelmében vizsgálja*

